



La nuova release di Apache Kafka disponibile in Cloudera Enterprise per uno streaming di dati sicuro e scalabile

PALO ALTO, Calif., 3 marzo, 2016 — [Cloudera](#), fornitore globale della piattaforma di analisi e gestione dei dati più veloce, semplice e sicura basata su [Apache Hadoop](#) e delle più recenti tecnologie open source, ha annunciato oggi l'ultima major release di [Apache Kafka](#), un sistema di messaggistica pubblica-sottoscrive fault tolerant ed estremamente scalabile creato per flussi di dati in tempo reale. Integrato all'interno di Cloudera Enterprise, questa versione più recente di Kafka offre funzioni di sicurezza critiche, progressi nelle operazioni multi-tenant e un'esperienza di sviluppo semplificata per pipeline di big data, permettendo agli utenti di assorbire e sfruttare più facilmente il valore dei crescenti volumi di streaming di dati che derivano dall'enorme numero di dispositivi connessi.

McKinsey prevede che entro il 2020, circa 30 miliardi di dispositivi saranno collegati all'Internet of Things (IoT) e il 40% del valore correlato potrà essere sbloccato abilitando l'interoperabilità e combinando i dati provenienti da molteplici fonti e sistemi IoT. Le aziende di qualsiasi settore si stanno rendendo conto del potenziale di questi dati di flusso in tempo reale nel generare nuovo valore e un vantaggio competitivo. Sia che si tratti della possibilità di accedere ai dati in tempo reale in combinazione con i dati storici per fornire approfondimenti che stimolino il coinvolgimento del cliente, migliorino in modo significativo la cura dei pazienti, potenzino il rilevamento delle frodi o la prossima generazione di aziende digitali, è difficile ignorare il ruolo che i dati di flusso giocano all'interno di un'azienda moderna. Kafka è studiato per far sì che le aziende possano sfruttare questi dati a qualsiasi livello.

[Cisco WebEx](#), leader nei servizi di conferenza e cliente di Cloudera, ha migliorato le valutazioni dei clienti e individuato una quantità di frodi superiore di 17 volte dal momento in cui è passata da diversi silos di dati compartimentati a un ambiente unificato per la discovery e l'analisi dei dati basato su Cloudera Enterprise e Server Cisco UCS. WebEx elabora i dati di flusso in tempo reale tramite Apache Spark e condivide i dati con i suoi team dedicati ai servizi e alle frodi tramite Kafka in modo che vengano avvisati di qualsiasi anomalia operativa mentre sono in corso le conferenze e possano agire o risolvere i problemi con tempestività.

Inoltre, [Cerner](#) ha sviluppato soluzioni per il monitoraggio dei pazienti che riuniscono molteplici fonti di dati sanitari per rilevare pericolose infezioni del sangue che richiedono attenzione immediata, salvando centinaia di vite dei pazienti. [Bidtellect](#) garantisce che gli inserzionisti di tutto il mondo traggano beneficio dalla stessa intelligence utilizzando Kafka per effettuare lo streaming dei dati da molteplici sedi verso l'hub di dati aziendali di Cloudera. Infine, [Cox Automotive](#) ha sviluppato dashboard in tempo reale per monitorare le principali applicazioni e misurazioni IT utilizzando Spark Streaming e Kafka.

“Disporre di pipeline sicure e affidabili per i dati in tempo reale non è mai stato così importante. I nostri clienti in ogni settore si trovano ad affrontare una sfida enorme: l'acquisizione e l'inserimento di enormi quantità di dati derivanti dalla crescente ondata di dispositivi IoT, in particolare nel momento in cui cercano di proteggere e gestire questi dati mentre fluiscono all'interno dei loro data hub aziendali” spiega Charles Zedlewski, vice presidente, Products di Cloudera. “Ora che l'ultima versione di Kafka è integrata direttamente nella piattaforma di Cloudera, i nostri clienti possono garantire che le loro pipeline di dati soddisfano gli stessi severi requisiti di sicurezza adottati dal resto dell'azienda. Grazie alle funzionalità enterprise aggiuntive, come i riavvii graduali e funzionalità di monitoraggio e troubleshooting all'avanguardia, i clienti possono concentrarsi sul valore che queste nuove fonti di dati e applicazioni forniscono, e non sulla gestione manuale degli strumenti basilari”.

Per facilitare l'inserimento di dati in tempo reale su larga scala e l'uscita all'interno di ambienti di produzione Hadoop, Cloudera ha incorporato l'ultima release di Kafka nella sua distribuzione per far sì che gli utenti godano di streaming sicuro, multi-tenancy affidabile e di un ambiente semplificato. Unitamente alle funzionalità di Cloudera Enterprise 5.5, questa versione aggiunge una disponibilità ininterrotta e funzionalità di monitoraggio e troubleshooting più efficaci, oltre a connessioni con i principali strumenti di terze parti per l'elaborazione dei flussi e l'integrazione dei dati. Qui di seguito le funzionalità specifiche migliorate:

- **Sicurezza efficace:** la crittografia via cavo end-to-end protegge i dati che circolano all'interno del sistema e tra i confini del centro dati, mentre l'autenticazione Kerberos previene gli accessi non autorizzati con un programma di gestione delle identità standard unificato che abbraccia l'intera piattaforma.
- **Multi-Tenancy affidabile:** limita i singoli client o tenant sulla base dei vincoli legati alle risorse per scalare in modo affidabile e supportare i crescenti volumi e fonti di dati, senza compromettere gli altri utenti.

- **Gestione aziendale:** Cloudera Manager permette di distribuire in modo semplice Kafka come parte di Cloudera Enterprise, con monitoraggio e avvisi personalizzati come parte della piattaforma. Con i riavvii graduali per Kafka, policy per la riproduzione configurabili e attività di troubleshooting estremamente veloci, Cloudera offre disponibilità costante per Kafka e le pipeline che vi fanno affidamento.
- **Semplici pipeline end-to-end:** una nuova API Java migliora l'esperienza degli sviluppatori nel collegare Kafka al resto dell'ecosistema di big data, inclusi strumenti come Apache Flume e Spark.

Sfruttando Kafka come parte di un abbonamento Cloudera Enterprise, i data engineer possono far fluire, elaborare e fornire i dati in tempo reale, tutto all'interno di un'unica piattaforma unificata. Con l'accesso al più ampio ecosistema di partner di Cloudera e un rigoroso programma di certificazione delle terze parti, questi utenti possono estendere le funzionalità della piattaforma con integrazioni dirette e affidabili con i principali strumenti per l'integrazione e l'arricchimento dei dati come Pentaho, Streamsets, Syncsort e Talend.

"L'integrazione diretta di questa nuova release di Kafka all'interno di Cloudera Enterprise è una notizia fantastica per la community di Kafka perché garantisce che Kafka è supportato da un importante fornitore di distribuzioni Hadoop ed è parte di una piattaforma di dati moderna pronta per l'ambiente di produzione", ha affermato Eddie White, EVP of business development di Pentaho, un'azienda di Hitachi Group. "In qualità di partner strategico anche Pentaho ne trae beneficio perché possiamo portare Kafka all'interno dei Pentaho Labs tramite la piattaforma Cloudera per convalidare ulteriormente la tecnologia con la nostra piattaforma per l'integrazione e l'analisi dei dati per aiutare le aziende ad accelerare i loro investimenti in ambito IoT, ora e in futuro".

"In StreamSets siamo lieti del costante sviluppo di Apache Kafka quale componente integrale di Cloudera Enterprise," spiega Arvind Prabhakar, chief technology officer, StreamSets. "StreamSets Data Collector, che si installa come pacchetto di Cloudera Manager, combina il design visivo della pipeline e le funzionalità di monitoraggio intelligente dei dati con Kafka per permettere ai clienti di Cloudera di distribuire i flussi di dati attraverso la piattaforma con una semplicità di sviluppo e una visibilità operativa senza precedenti".

"I nostri clienti nell'ambito dei servizi finanziari, sanità, retail e telecomunicazioni cercano di sfruttare la velocità e la flessibilità di Apache Kafka per servizi dati fault tolerant e a bassa latenza per un crescente numero di casi d'uso come il rilevamento

delle frodi e l'analisi dei dati relativi alla sicurezza e alla telemetria", ha dichiarato Tendü Yoğurtçu, General Manager della divisione Big Data di Syncsort. "Cloudera supporta tale necessità fornendo tempestivamente la nuova versione di Kafka in Cloudera Enterprise, proteggendo le pipeline di dati in tempo reale. L'integrazione di Syncsort con Kafka aiuta le aziende a utilizzare un unico ambiente software semplice da usare per creare una pipeline di dati per fonti aziendali diverse, tra cui dati batch, streaming, mainframe e IoT".

Secondo Ashley Stirrup, CMO di Talend: "Apache Kafka si sta rapidamente affermando come importante protocollo di messaggistica per scenari big data in tempo reale. Con questa nuova release da parte di Cloudera, i nostri clienti congiunti possono creare applicazioni intelligenti per smart city, raccomandazioni in tempo reale, manutenzione predittiva, gaming di precisione e molto altro. Combinando Kafka per le attività di inserimento e Apache Spark per l'elaborazione dei dati e machine learning, i clienti possono sfruttare approfondimenti di dettaglio 'puntuali' come mai prima".

Kafka è cresciuto fino a diventare una tecnologia open source affidabile per acquisire e trasportare flussi di dati ad alta velocità da dispositivi e sensori Internet of Things. Con l'aiuto dell'ecosistema di Cloudera con oltre 2.000 partner e tecnologie all'avanguardia per la gestione dei dati enterprise, Kafka è oggi meglio equipaggiato per aiutare le aziende a sfruttare le opportunità offerte dall'IoT su larga scala.

###

A proposito di Cloudera

Cloudera offre la piattaforma per la gestione e l'analisi dei dati aziendali basata su Apache Hadoop e le più recenti tecnologie di open source. Le principali aziende di tutto il mondo si affidano a Cloudera per affrontare le principali problematiche di business con Cloudera Enterprise, la piattaforma dati più veloce, semplice e sicura disponibile ad oggi. I clienti di Cloudera sono in grado di acquisire, archiviare, elaborare e analizzare efficientemente una vasta quantità di dati per utilizzare analisi avanzate al fine di elaborare decisioni di business più velocemente, in modo più flessibile e con costi inferiori rispetto al passato. Per assicurare il successo dei propri clienti, Cloudera offre ampio supporto, sessioni di training e servizi professionali. Ulteriori informazioni sono disponibile su <http://cloudera.com>.

Per saperne di più:

Web: cloudera.com

Blog: cloudera.com/blog e vision.cloudera.com

Twitter: twitter.com/clouderaITA

Facebook: facebook.com/cloudera

Linkedin Group: [linkedin.com/Cloudera Italia](https://linkedin.com/Cloudera%20Italia)

Cloudera Community: community.cloudera.com